AD-A042 729 WISCONSIN UNIV MADISON MATHEMATICS RESEARCH CENTER F/G 12/1
THE CONTINUOUS TIME BAYES' SEQUENTIAL PROCEDURE FOR ESTIMATING --ETC(U)
JUN 77 C P SHAPIRO, R WARDROP DAAG29-75-C-0024
UNCLASSIFIED MRC-TSR-1753 NL

1 OF 1
ADA042 729

END
DATE
FILMED
9-77
DDC

12

MRC Technical Summary Report #1753

THE CONTINUOUS TIME BAYES' SEQUENTIAL PROCEDURE FOR ESTIMATING THE ARRIVAL RATE OF A POISSON PROCESS AND LARGE SAMPLE PROPERTIES

C. P. Shapiro and Robert Wardrop

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

June 1977

(Received May 5, 1977)

Sponsored by

U.S. Army Research Office
P.O. Box 12211
Research Triangle Park
North Carolina 27709

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

THE CONTINUOUS TIME BAYES' SEQUENTIAL PROCEDURE FOR ESTIMATING THE ARRIVAL RATE OF A POISSON PROCESS AND LARGE SAMPLE PROPERTIES

C. P. Shapiro and Robert Wardrop

Technical Summary Report #1753
June 1977

ABSTRACT

Let $X(t)$, $t \geq 0$, be a homogeneous Poisson process with arrival rate θ. Sequential estimation procedures $(\sigma, \hat{\theta}_\sigma)$ are considered with loss due to estimation $L(\theta,\hat{\theta}) = \theta^{-1}(\theta-\hat{\theta})^2$, and sampling costs involving both time and arrival costs. In this context the Bayes' sequential procedure is obtained in a simple computable form. The large sample properties of the procedure are then studied when θ is fixed but unknown, and the Bayes' stopping rule τ is shown to be asymptotically equivalent to the best fixed sample size procedure when θ is known. Asymptotic normality of the Bayes' sequential estimator $\hat{\theta}_\tau$ of θ is also shown.

AMS(MOS) Subject Classification - 62L12, 62C10

Key Words - Poisson process, characteristic operator, Dynkin's identity, sequential Bayes' estimator.

Work Unit Number 4 - Probability, Statistics and Combinatorics

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024.

THE CONTINUOUS TIME BAYES' SEQUENTIAL PROCEDURE
FOR ESTIMATING THE ARRIVAL RATE OF A
POISSON PROCESS AND LARGE SAMPLE PROPERTIES

C. P. Shapiro and Robert Wardrop

1. Introduction. Conditional on the value of θ, $\theta > 0$, let $X(t)$, $t \geq 0$, be a homogeneous Poisson process with rate θ. Let $\mathfrak{F}(t)$, $t \geq 0$, denote the sigma algebra generated by $X(s)$, $0 \leq s \leq t$. Sequential estimation procedures of the form $(\sigma, \hat{\theta}_\sigma)$ are studied, where σ is a stopping time with respect to $\mathfrak{F}(t)$, and $\hat{\theta}_\sigma$ is an $\mathfrak{F}(\sigma)$ measurable random variable, with $\mathfrak{F}(\sigma)$ the sigma algebra of events prior to σ.

The loss due to estimation is

$$L(\theta,\hat{\theta}) = \theta^{-1}(\theta-\hat{\theta})^2 \quad . \tag{1.1}$$

This loss measures estimation error in terms of standard deviation θ^{-1} forcing more precision at smaller θ-values. The loss was suggested by Dvoretzky, Kiefer, and Wolfowitz (1953) and Hodges and Lehmann (1951).

The cost of sampling consists of two components: $c_A > 0$, the cost of observing one arrival, and $c_T > 0$, the cost of observing the process for one unit of time.

In Section 2, the Bayes sequential procedure, $(\tau, \hat{\theta}_\tau)$, is derived using a gamma prior on θ and the loss and cost structures described above. In Sections 3 and 4, the asymptotic properties of $(\tau, \hat{\theta}_\tau)$ are examined without reference to the Bayesian origin of the procedure. The limiting form of τ is given in Theorem 3.1 and the asymptotic normality of $\hat{\theta}_\tau$ is given in Theorem 4.1. Concluding remarks along with methods for choosing c_A and c_T are given in Section 5.

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024.

2. The Bayes' sequential procedure. For a given prior distribution on θ, the Bayes' sequential procedure (BSP) minimizes the total expected cost,

$$E(\theta^{-1}(\theta-\hat{\theta}_\sigma)^2) + E(c_A X(\sigma) + c_T\sigma) \tag{2.1}$$

over all pairs $(\sigma, \hat{\theta}_\sigma)$. For a fixed stopping time σ, the first term in (2.1) can be written as $E(E(\theta^{-1}(\theta-\hat{\theta}_\sigma)^2 | \mathfrak{F}(\sigma)))$ and is minimized by taking $\hat{\theta}_\sigma$ to be the Bayes' estimator of θ given $\mathfrak{F}(\sigma)$. Hence, the first step in finding the BSP is to determine the posterior distribution of θ given $\mathfrak{F}(\sigma)$, for any σ.

Henceforth, let θ have prior density $\pi(\theta) = \Gamma(\alpha_0)^{-1} \beta_0^{\alpha_0}\theta^{\alpha_0-1} e^{-\theta\beta_0}$ for $\theta > 0$, $\alpha_0 > 1$, and $\beta_0 > 0$, abbreviated gamma (α_0, β_0). With this prior density, the posterior distribution of θ given $\mathfrak{F}(t)$, is gamma (α_t, β_t), where $\alpha_t = \alpha_0 + X(t)$ and $\beta_t = \beta_0 + t$. Using the loss defined in (1.1), the Bayes' estimator of θ given $\mathfrak{F}(t)$ is $\hat{\theta}_t = \beta_t^{-1}(\alpha_t - 1)$ and the posterior expected loss using $\hat{\theta}_t$ is β_t^{-1}. For σ a stopping time, the posterior distribution of θ given $\mathfrak{F}(\sigma)$ is gamma $(\alpha_\sigma, \beta_\sigma)$ by the strong Markov property of the Poisson process. Thus, the Bayes' estimator of θ given $\mathfrak{F}(\sigma)$ is $\hat{\theta}_\sigma = \beta_\sigma^{-1}(\alpha_\sigma - 1)$ with posterior expected loss $E(L(\theta, \hat{\theta}_\sigma) | \mathfrak{F}(\sigma)) = \beta_\sigma^{-1}$. For $t \geq 0$, define

$$\mathcal{C}_t = \beta_t^{-1} + c_A X(t) + c_T t \,. \tag{2.2}$$

The total cost of the procedure $(\sigma, \hat{\theta}_\sigma)$ is $\mathcal{C}_\sigma$, and the expected total cost is $E(\mathcal{C}_\sigma)$. The Bayes sequential procedure minimizes $E(\mathcal{C}_\sigma)$.

Define the stopping rule τ by

$$\tau = \text{least } t \geq 0 \text{ such that } c_A\alpha_t + c_T\beta_t \geq \beta_t^{-1} \,. \tag{2.3}$$

Note that $P(\tau < \infty) = 1$, and that τ is very simple to use. Rule τ will be shown to be Bayes' in Theorem 2.1. Before motivating the rule, Lemma 2.1 gives bounds for τ and $X(\tau)$.

Lemma 2.1. For τ given by (2.3),

i) $\tau \leq \min(c_A^{-1}\alpha_0^{-1}, c_T^{-1/2})$

ii) $X(\tau) \leq (c_A\beta_0)^{-1} + 1$.

Proof: Fix $\varepsilon > 0$. Then if $\tau > \varepsilon$, $\tau - \varepsilon$ must satisfy the reverse inequality in

expression (2.3) defining τ . Thus,

$$c_T \beta_{\tau-\varepsilon} + c_A \alpha_{\tau-\varepsilon} < \beta_{\tau-\varepsilon}^{-1} \; .$$

This inequality first implies that $c_T \beta_{\tau-\varepsilon} < \beta_{\tau-\varepsilon}^{-1}$ which gives $\tau - \varepsilon < c_T^{-1/2}$. Secondly, it implies that $c_A \alpha_{\tau-\varepsilon} < \beta_{\tau-\varepsilon}^{-1}$ which gives $c_A \beta_{\tau-\varepsilon} < (\alpha_{\tau-\varepsilon})^{-1} \leq \alpha_0^{-1}$ and thus, $\tau - \varepsilon < (\alpha_0 c_A)^{-1}$. Also this second form implies $c_A \alpha_{\tau-\varepsilon} < \beta_0^{-1}$ which gives $c_A X(\tau-\varepsilon) < \beta_0^{-1}$ and hence $X(\tau-\varepsilon) < (c_A \beta_0)^{-1}$. Since ε is arbitrary the proof is complete.

The motivation for rule τ is given in Lemma 2.3. Briefly, the rule is derived from the characteristic operator of the marginal process, $X(t)$, obtained from the Poisson process by mixing over θ according to the prior distribution. Rules derived from characteristic operators have been considered by other authors, namely, Ross (1971) and Starr, Wardrop, and Woodroofe (1976), and are sometimes called "infinitesimal look-ahead rules." Such rules are essentially continuous time analogues of rules derived from the monotone case (Chow, Robbins, Siegmund, 1971).

Recall the definition of $\mathcal{C}_t$ in (2.2). Define

$$A\,\mathcal{C}_t = \lim_{h \downarrow 0} \frac{E(\mathcal{C}_{t+h} \mid (t, X(t)) - \mathcal{C}_t}{h} \tag{2.4}$$

when the limit exists. Likewise, for real valued bounded functions f on $[0,\infty)$, define

$$A_t\, f(x) = \lim_{h \downarrow 0} \frac{E(f(X(t+h) \mid (t, X(t)) = (t,x)) - f(x)}{h} \tag{2.5}$$

when the limit exists. The expression in (2.5) is the characteristic operator of the process X evaluated at the function f . Note that the marginal distribution of X (not conditional on θ) is used to define (2.4) and (2.5). Thus, the process in question is not Poisson, but is obtained from the Poisson by mixing over the parameter θ . Crucial in showing the optimality of τ is Lemma 2.2 which states that $X(t)$ is marginally a strong Markov process.

Lemma 2.2. Suppose that conditional on θ, $X(t)$ is a homogeneous Poisson process with arrival rate θ . Then for any prior distribution Λ on θ , $X(t)$ is marginally a strong Markov process.

Proof. Using the fact that $P(X(t_i) = k_i, \quad i = 1,2,\ldots,m)$ is equal to (for $t_i \le t_{i+1}$) $\int \prod_{i=1}^{m} P_\theta(X(t_i - t_{i-1}) = k_i - k_{i-1})\, d\Lambda(\theta)$, the Markov property follows from a straightforward computation. Since $X(t)$ is marginally a pure jump process, $X(t)$ is strong Markov.

Lemma 2.3. For all $t \ge 0$,

i) $$A\mathcal{C}_t = -\beta_t^{-2} + c_A \alpha_t \beta_t^{-1} + c_T ,$$

ii) $$A_t x = (\alpha_0 + x)(\beta_0 + t)^{-1} .$$

Proof: $E[\mathcal{C}_{t+h} - \mathcal{C}_t | \alpha_t, \theta]$

$$= [(\beta_t + h)^{-1} - \beta_t^{-1}] + c_A E[X(t+h) - X(t) | \alpha_t, \theta] + c_T h$$

$$= [(\beta_t + h)^{-1} - \beta_t^{-1}] + c_A \theta h + c_T h .$$

Thus, since $E(\theta | \alpha_t) = \alpha_t \beta_t^{-1}$, $E[\mathcal{C}_{t+h} - \mathcal{C}_t | \alpha_t]$ is equal to

$$[(\beta_t + h)^{-1} - \beta_t^{-1}] + c_A \alpha_t \beta_t^{-1} h + c_T h .$$

Dividing by h and taking the limit as h tends to zero gives the result. Note that (ii) is shown in the body of the proof.

From Lemma 2.3, rule τ can be expressed as

$$\tau = \text{least } t \ge 0 \text{ such that } A\mathcal{C}_t \ge 0 . \tag{2.6}$$

Also, note that $A\mathcal{C}_t$ changes sign only once.

The next lemma shows that Dynkin's identity (Brieman, p. 376) holds for $E\mathcal{C}_\sigma$ for a large class of stopping times σ. Note that $\mathcal{C}_0 = \beta_0^{-1}$.

Lemma 2.4. Suppose σ is a stopping time such that $E(\sigma) < \infty$ and that $X(\sigma)$ is bounded. Then $E\mathcal{C}_\sigma = \beta_0^{-1} + E\int_0^\sigma A\,\mathcal{C}_t\, dt$.

Proof. Note that $E\,\mathcal{C}_\sigma < \infty$ under the hypotheses of the lemma, and express $E\,\mathcal{C}_\sigma = E\beta_\sigma^{-1} + c_A EX(\sigma) + c_T E(\sigma)$. Lemma 2.3 gives $A\mathcal{C}_t = -\beta_t^{-2} + c_A \alpha_t \beta_t^{-1} + c_T$. Simple integration yields $E\beta_\sigma^{-1} = E\int_0^\sigma -\beta_t^{-2}\, dt + \beta_0^{-1}$ and $E(\sigma) = E\int_0^\sigma 1\, dt$. Thus, the proof is complete if $E\,X(\sigma)$ is shown to be equal to $E\int_0^\sigma \alpha_t \beta_t^{-1}\, dt$. A straightforward truncation argument is used to show this. Since $X(\sigma)$ is bounded there exists $M < \infty$ such that $X(\sigma) < M$. Fix $\varepsilon > 0$, and define $f^*(x)$ increasing and differentiable for all $x > 0$

such that $f^*(x) = x$ for $x \le M$ and $f^*(x) = M + \varepsilon$ for $x > M + \varepsilon$. Then f^* is bounded, $A_t f^*$ exists, and Dynkin's identity implies that $Ef^*(X(\sigma)) = E\int_0^\sigma A_t f^*(X(t))dt$. But $f^*(X(\sigma)) = X(\sigma)$ since $X(\sigma) < M$. Also, $f^*(X(t)) = X(t)$ for all $t \le \sigma$ since $X(t) \le X(\sigma)$ on $[t \le \sigma]$. Thus, $Ef^*(X(\sigma)) = EX(\sigma)$, and $A_t f^*(X(t)) = A_t(X(t)) = \alpha_t \beta_t^{-1}$ on $[t \le \sigma]$.

<u>Theorem 2.1</u>. For all stopping times σ, $E(\mathcal{C}_\tau) \le E(\mathcal{C}_\sigma)$, where τ is given by (2.3).

<u>Proof</u>. Lemma 2.1 implies $E(\tau) < \infty$ and $X(\tau)$ bounded. Thus,

$$E(\mathcal{C}_\tau) = E\int_0^\tau A\mathcal{C}_t\, dt + \beta_0^{-1} < \infty$$

by Lemma 2.4.

Let $S = \{\sigma: E(\sigma) < \infty$ and $X(\sigma)$ bounded$\}$. If σ is in S, then $E\mathcal{C}_\sigma = \beta_0^{-1} + E\int_0^\sigma A\mathcal{C}_t\, dt$. Hence, $E\mathcal{C}_\sigma - E\mathcal{C}_\tau = E\int_{[\sigma \ge \tau]} A\mathcal{C}_t\, dt - E\int_{[\sigma < \tau]} A\mathcal{C}_t\, dt$ which is nonnegative since $A\mathcal{C}_t \ge 0$ on $[\sigma \ge \tau]$ and $A\mathcal{C}_t < 0$ on $[\sigma < \tau]$. Thus, τ is optimal in S.

If $E\mathcal{C}_\sigma = \infty$, then τ is obviously better. Thus consider σ, $E\mathcal{C}_\sigma < \infty$, and choose a sequence of integers m_k increasing to ∞. Define stopping rule $\sigma_k = \sigma$ if $X(\sigma) \le m_k$ and $= t_k$ if $X(\sigma) > m_k$, where t_k is the smallest t such that $X(t) = m_k$. Then $X(\sigma_k) \le m_k$ and $E\sigma_k \le E\sigma < \infty$, and hence σ_k is in S for all k. Since τ is optimal in S, $E\mathcal{C}_\tau \le E\mathcal{C}_{\sigma_k}$ for all k. Thus, the proof of the theorem will be complete if $E\mathcal{C}_{\sigma_k}$ tends to $E\mathcal{C}_\sigma$ as k tends to ∞. Write $E\mathcal{C}_{\sigma_k} = E\mathcal{C}_\sigma[X(\sigma) \le m_k] + E\mathcal{C}_{t_k}[X(\sigma) > m_k]$. The first term tends to $E\mathcal{C}_\sigma$ by the monotone convergence theorem. For the second term, $E\mathcal{C}_{t_k}[X(\sigma) > m_k] \le \beta_0^{-1} P(X(\sigma) > m_k) + E\mathcal{C}_\sigma[X(\sigma) > m_k]$ since $X(\sigma) > m_k$ implies $t_k < \sigma$. Both terms above tend to 0 since $E\mathcal{C}_\sigma < \infty$.

3. Large sample properties of τ. In this section the Bayes' procedure $(\tau, \hat{\theta}_\tau)$ is examined in the classical framework. The parameter θ is considered fixed but unknown and all probabilities and expectations are conditional on θ and denoted by P_θ and E_θ, respectively. The procedure $(\tau, \hat{\theta}_\tau)$ does not minimize $E_\theta C_\sigma$ for all θ, but only the average of $E_\theta C_\sigma$ over the prior distribution of Section 2.

The large sample properties of the procedure are studied by letting the sampling costs tend jointly to zero. Define

$$t^* = t^*(\theta) = (c_A\theta + c_T)^{-1/2} . \tag{3.1}$$

The main result of this section (Theorem 3.1) is that τ is asymptotically equivalent to t^* as $\underline{c} = (c_A, c_T)$ tends to $\underline{0} = (0,0)$ with $c_A c_T^{-1}$ converging to $c_0 \leq \infty$.

As motivation for this limiting form of τ, note that $E_\theta C_t$ is equal to $\beta_t^{-1} + c_A\theta\, t + c_T t$. Let $H(x) = (\beta_0 + x)^{-1} + c_A\theta\, x + c_T x$. Then H attains a unique minimum at $x = (c_A\theta + c_T)^{-1/2} - \beta_0$. Ignoring β_0 gives t^*.

The following lemmas give rates and uniform integrability results needed in the proof of Theorem 3.1.

Lemma 3.1. For each $\varepsilon > 0$,

$$P_\theta(|(\tau/t^*) - 1| > \varepsilon) \leq 2\exp(c_A^{-1/2} D(\underline{c}, \varepsilon)) ,$$

where $D(\underline{c},\varepsilon) \to D(\varepsilon) < 0$ and finite as c_A, $c_T \to 0$ along any sequence for which $c_A c_T^{-1} \to c_0$, $0 < c_0 \leq \infty$.

Proof. Note that $\{\tau > t\} = \{X(t) < B_t\}$, where $B_t = (c_A\beta_t)^{-1} - c_T\beta_t c_A^{-1} - \alpha_0$. Let $t_\varepsilon = (1+\varepsilon)t^*$. Then $P_\theta(\tau/t^* > 1+\varepsilon) = P_\theta(X(t_\varepsilon) < B_{t_\varepsilon})$. Since $X(t_\varepsilon)$ is Poisson (θt_ε), Bernstein's inequality implies that for all $u > 0$,

$$\begin{aligned} P_\theta(X(t_\varepsilon) < B_{t_\varepsilon}) &\leq \exp(u\, B_{t_\varepsilon})\, E_\theta \exp(-u\, X(t_\varepsilon)) \\ &= \exp(u\, B_{t_\varepsilon} + \theta\, t_\varepsilon(e^{-u} - 1)) \\ &= \exp(c_A^{-1/2}\, u\, B(\underline{c}, t_\varepsilon, u)), \text{ where} \end{aligned}$$

$$B(\underline{c}, t_\varepsilon, u) = c_A^{-1/2}\beta_{t_\varepsilon}^{-1} - c_T c_A^{-1/2}\beta_{t_\varepsilon} - \alpha_0 c_A^{1/2} - \theta t_\varepsilon c_A^{-1/2} u^{-1}(1-e^{-u}) .$$

Since $c_A c_T^{-1} \to c_0 > 0$, it is easy to show that $c_A^{1/2} t_\varepsilon$ and $c_A^{1/2}\beta_{t_\varepsilon}$ both tend to

$(1 + \varepsilon)(\theta + c_0^{-1})^{-1/2}$. Thus, as sampling costs tend to 0, $B(\underline{c}, t_\varepsilon, u) \to B(\varepsilon, u)$, where

$$B(\varepsilon,u) = (1+\varepsilon)^{-1}(\theta + c_0^{-1})^{1/2} - (1+\varepsilon)(\theta + c_0^{-1})^{-1/2}(\theta u^{-1}(1 - e^{-u}) + c_0^{-1}) .$$

Now, the limit of $B(\varepsilon,u)$ as u tends to 0 is negative and finite, and hence $B(\varepsilon,u) < 0$ for all $u \le u_0$ depending on ε. Choosing $D^+(\underline{c},\varepsilon) = u_0 B(\underline{c}, t_\varepsilon, u_0)$ yields the appropriate rate for $P_\theta(\tau/t^* > 1 + \varepsilon)$. For the lower probability, define $t_\varepsilon^- = (1 - \varepsilon)t^*$ and write $P_\theta(\tau/t^* < 1 - \varepsilon) = P_\theta(X(t_\varepsilon^-) \le B_{t_\varepsilon^-}) \le \exp(uc_A^{-1/2}(-B(\underline{c}, t_\varepsilon^-, -u)))$. As above, there exists $u_0 > 0$ such that $B(\underline{c}, t_\varepsilon^-, -u_0) > 0$, and thus $D^-(\underline{c},\varepsilon)$ may be chosen as $-u_0 B(\underline{c}, t_\varepsilon^-, -u_0)$. The proof is completed by taking $D(\underline{c}, \varepsilon) = \max(D^+, D^-)$.

<u>Lemma 3.2</u>. For any $\varepsilon > 0$,

$$P_\theta(|(\tau/t^*) - 1| > \varepsilon) \le 2\exp(c_T^{1/2}\, c_A^{-1}\, D(\underline{c},\varepsilon)) ,$$

where $D(\underline{c},\varepsilon) \to D(\varepsilon) < 0$ and finite as $c_A, c_T \to 0$ along any sequence such that $c_A\, c_T^{-1} \to 0$.

<u>Proof</u>. As in the proof of Lemma 3.1,

$$\begin{aligned} P_\theta(\tau/t^* > 1 + \varepsilon) &\le \exp(u\, B_{t_\varepsilon} - \theta t_\varepsilon(1 - e^{-u})) \\ &= \exp(c_T^{1/2} c_A^{-1}\, u\, G(\underline{c}, t_\varepsilon, u)) . \end{aligned}$$

Using the fact that $c_T^{1/2}\, t_\varepsilon$ and $c_T^{1/2}\, \beta_{t_\varepsilon}$ both tend to $(1+\varepsilon)$ and the methods of Lemma 3.1 complete the proof.

<u>Lemma 3.3</u>. As $c_A, c_T \to 0$ such that $c_A\, c_T^{-1} \to c_0 \le \infty$,

i) τ/t^* is uniformly integrable (P_θ),

ii) $t^*/(\beta_0 + \tau)$ is uniformly integrable (P_θ).

<u>Proof</u>. Let $Y = Y_{\underline{c}} = |(\tau/t^*) - 1|$. It suffices to show that for $a > 0$ fixed, $\lim_{\underline{c}\to 0} \int_{\{Y>a\}} Y \, dP_\theta = 0$ as $c_A c_T^{-1} \to c_0 \le \infty$. First suppose that $c_A c_T^{-1} \to c_0 > 0$. From Lemma 2.1, $Y \le \alpha_0^{-1}\, c_A^{-1/2}(\theta + c_T c_A^{-1})^{1/2}$. Thus, Lemma 3.1 implies that

$$\int_{\{Y>a\}} Y \, dP_\theta \le \alpha_0^{-1}\, c_A^{-1/2}(\theta + c_T c_A^{-1})^{1/2} \exp(c_A^{-1/2}\, D(\underline{c}, a))$$

which tends to 0 as $c_A c_T^{-1} \to c_0 > 0$. Secondly, suppose $c_A c_T^{-1} \to 0$. Then Lemma 2.1 implies $Y \leq (c_A c_T^{-1} \theta + 1)^{1/2}$. Thus, Lemma 3.1 implies that

$$\int_{\{Y>a\}} Y \, dP_\theta \leq (c_A c_T^{-1} \theta + 1) \exp(c_T^{1/2} c_A^{-1} D(\underline{c},a))$$

which tends to 0.

To prove (ii), note that $t^*/(\beta_0 + \tau) \leq t^*/\beta_0 = (c_A\theta + c_T)^{-1/2} \beta_0^{-1}$, and apply the same techniques as in (i).

<u>Theorem 3.1</u>. If $c_A, c_T \to 0$ such that $c_A c_T^{-1} \to c_0 \leq \infty$, then

i) $\tau/t^* \to 1$ (in P_θ probability),

ii) $E_\theta(\tau/t^*) \to 1$.

<u>Proof</u>. The results follow immediately from Lemmas 3.1-3.3.

<u>Corollary 3.1</u>. If $c_A, c_T \to 0$ such that $c_A c_T^{-1} \to c_0 \leq \infty$, then $t^* E_\theta \mathcal{C}_\tau \to 2$.

<u>Proof</u>. $E_\theta \mathcal{C}_\tau = E_\theta(\beta_0 + \tau)^{-1} + c_A\theta E(\tau) + c_T E(\tau)$ by Wald's Lemma. Thus, $t^* E_\theta \mathcal{C}_\tau = E_\theta(t^*/(\beta_0+\tau)) + E_\theta(\tau/t^*) \to 1 + 1 = 2$ by Theorem 3.1 and Lemma 3.3.

4. Asymptotic normality of $\hat{\theta}_\tau$. Once the limiting form of τ is found, the asymptotic distribution of $\hat{\theta}_\tau$ is obtained by standard methods. First a well known result for random sums of random variables is stated.

Lemma 4.1. Suppose $Y_1, Y_2, \ldots$ are independent and identically distributed with mean 0 and variance 1, and that N is an integer valued random variable tending to ∞ as sampling costs tend to 0. If there exists n^*, nonnrandom, such N/n^* tends to 1 (in probability) then

$$(N)^{-1/2} \sum_{i=1}^{N} Y_i \to Z \quad \text{(in distribution)} ,$$

where Z is normal with mean zero and variance 1.

Proof. See Renyi (1957).

Lemma 4.2. As sampling costs c_A, c_T tend jointly to zero along any sequence such that $c_A c_T^{-1} \to c_0 \leq \infty$,

$$(\tau\theta)^{-1/2} (X(\tau) - \tau\theta) \to Z \quad \text{(in distribution)}$$

where Z is normal with mean zero and variance one.

Proof. For each $n = 1,2,\ldots$, $X(n) = \sum_{i=1}^{n} Y_i$, with $Y_i = X(i) - X(i-1)$ independent and identically distributed Poisson (θ). Define $N = [\tau + 1]$, where $[.]$ is the greatest integer function. Then N is an integer valued random variable with respect to $\{\mathfrak{F}_n\}$, with $\mathfrak{F}_n$ the sigma algebra generated by $Y_1,\ldots,Y_n$. Also, Theorem 3.1 implies that N/t^* tends to 1 in probability since $\tau < N \leq \tau + 1$. Thus, Lemma 4.1 implies that $(N\theta)^{-1/2}(X(N) - \theta N)$ tends to Z in distribution.

Now, $\tau < N \leq \tau + 1$ implies $X(N) \geq X(\tau)$ and $X(N) - X(\tau) \leq X(\tau + 1) - X(\tau)$. Thus, for any $\varepsilon > 0$,

$$P_\theta((t^*)^{-1/2} (X(N) - X(\tau)) > \varepsilon) \leq (t^*)^{-1/2} \varepsilon^{-1} E_\theta(X(\tau+1) - X(\tau))$$

$= (t^*)^{-1/2} \varepsilon^{-1} \theta$ which tends to 0. Thus, $(t^*)^{-1/2} X(N) - (t^*)^{-1/2} X(\tau)$ tends to 0 in probability. Next, since N/t^* tends to 1 in probability, then $N^{-1/2} X(N) - N^{-1/2} X(\tau)$ tends to 0 in probability. Thus,

$(N\theta)^{-1/2}(X(\tau) - N\theta)$ tends to Z in distribution. Using N/τ tends to 1 in probability gives $(\tau\theta)^{-1/2}(X(\tau) - N\theta)$ tending to Z in distribution. Finally,

$$(\tau\theta)^{-1/2}(X(\tau) - \tau\theta) = (\tau\theta)^{-1/2}(X(\tau) - N\theta) + (\tau\theta)^{-1/2}(\theta N - \theta\tau) \quad ,$$

where the last term tends to 0 in probability since $N - \tau \leq 1$.

Theorem 4.1. $(\tau)^{1/2} \dfrac{(\hat{\theta}_\tau - \theta)}{\sqrt{\theta}} \to Z$ (in distribution) where Z is normal with mean 0 and variance 1 .

Proof. $\dfrac{(\tau)^{1/2}}{\sqrt{\theta}}(\hat{\theta}_\tau - \theta)$

$$= \frac{\tau}{\beta_0 + \tau}\{(\tau\theta)^{-1/2}(X(\tau) - \theta\tau)\} + \frac{\sqrt{\tau}}{\sqrt{\theta}\,(\beta_0 + \tau)}\{\alpha_0 - 1 - \theta\beta_0\} \quad .$$

The second term tends to 0 in probability. The 1st term tends to Z in distribution using Lemma 4.2.

5. Concluding remarks. Although Bayesian methods are used to derive the procedure $(\tau, \hat{\theta}_\tau)$, the procedure has desirable properties in the classical framework, and these properties are of course independent of the prior distribution of Section 2. In particular, the rule τ is shown to be asymptotically equivalent to t^* (Theorem 3.1) where t^* is approximately the best fixed sample size procedure when θ is known. Thus, the asymptotic equivalence of τ and t^* is a very strong property.

The stopping rule τ also has desirable small sample properties. A common criticism of sequential procedures is that they may be unbounded. As shown in Lemma 2.2, τ is bounded, and in addition, $X(\tau)$ is bounded. Since in applications, the arrivals of a Poisson process may represent the failures of expensive experimental units, or failures of treatments, a rule σ with $X(\sigma)$ bounded may be desirable for monetary or ethical considerations.

Furthermore, the bounds given in Lemma 2.2 may guide the experimenter in his choice of costs. If the experimenter decides that the experiment must terminate by time A or that the number of arrivals observed must not exceed B, then costs c_A and c_T can be selected so that the bounds given in Lemma 2.1 are less than A and B respectively.

REFERENCES

[1] Brieman, L. (1968). Probability. Addison-Wesley, Reading.

[2] Chow, Y. S., Robbins, H., and Siegmund, D. (1971). Great Expectations: The Theory of Optimal Stopping. Houghton-Mifflin, Boston.

[3] Dvoretzky, A., Kiefer, J., and Wolfowitz, J. (1953). Sequential decision problems for processes with continuous time parameter. Problems of estimation. Ann. Math. Statist. 24, 403-415.

[4] Hodges, J. L. and Lehmann, E. L. (1951). Some applications of the Cramér-Rao inequality. Proc. Second Berkeley Symp. Math. Statist. Prob., Univ. of California Press.

[5] Renyi, A. (1957). On the asymptotic distribution of the sum of a random number of independent random variables. Acta. Math. 8, 193-199.

[6] Ross, S. M. (1971). Infinitesimal look-ahead stopping rules. Ann Math. Statist. 42, 297-303.

[7] Starr, N., Wardrop, R., and Woodroofe, M. (1976). Estimating a mean from delayed observations. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 35, 103-113.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER #1753	2. GOVT ACCESSION NO. (9)	3. RECIPIENT'S CATALOG NUMBER Technical
4. TITLE (and Subtitle) (6) THE CONTINUOUS TIME BAYES' SEQUENTIAL PROCEDURE FOR ESTIMATING THE ARRIVAL RATE OF A POISSON PROCESS AND LARGE SAMPLE PROPERTIES.		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) (10) C. P. Shapiro ~~and~~ Robert Wardrop		8. CONTRACT OR GRANT NUMBER(s) (15) DAAG29-75-C-0024
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin 610 Walnut Street Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Work Unit Number 4 - Probability, Statistics and Combinatorics
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office P.O. Box 12211 Research Triangle Park, North Carolina 27709	(11)	12. REPORT DATE June 1977
		13. NUMBER OF PAGES 12
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) (12) 15p.		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

(14) MRC-TSR-1753

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Poisson process, characteristic operator, Dynkin's identity, sequential Bayes' estimator

sigma, theta bar sub sigma

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Let $X(t)$, $t \geq 0$, be a homogeneous Poisson process with arrival rate θ. Sequential estimation procedures $(\sigma, \hat{\theta}_\sigma)$ are considered with loss due to estimation of $L(\theta, \hat{\theta}) = \theta^{-1}(\theta - \hat{\theta})^2$, and sampling costs involving both time and arrival costs. In this context the Bayes' sequential procedure is obtained in a simple computable form. The large sample properties of the procedure are then studied when θ is fixed but unknown, and the Bayes' stopping rule τ is shown to be asymptotically equivalent to the best fixed sample size procedure when θ is known. Asymptotic normality of the Bayes' sequential estimator $\hat{\theta}_\tau$ of θ is also shown.

squared tau theta bar sub tau theta

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

theta theta bar

221200

VB